DOCUMENT RESUME

ED 402 319 TM 025 811

AUTHOR Green, Kathy E.; Hutchinson, Susan R.
TITLE Reviewing the Research on Mail Survey Response Rates: Meta-Analysis.
PUB DATE Apr 96
NOTE 23p.; Paper presented at the Annual Meeting of the American Educational Research Association (New York, NY, April 8-12, 1996). For related papers, see TM 025 810-812.
PUB TYPE Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Comparative Analysis; *Effect Size; Followup Studies; *Incentives; *Mail Surveys; *Meta Analysis; Research Methodology; *Response Rates (Questionnaires); Responses
IDENTIFIERS Descriptive Research

ABSTRACT

Meta-analysis, a way of cumulating the results of research studies, focuses on the magnitudes of effect sizes and seeks to explain effects through study characteristics. This meta-analysis used the methods developed by G. V. Glass to summarize the research on mail survey response rate manipulations. A narrative review using the same studies set the stage for this review and a related review using descriptive methods. Results will demonstrate the similarities and differences one may see in analyzing the same data in different ways, as well as indicating the effectiveness of treatments meant to enhance mail survey response rates. Studies were coded and repeated measure main effects were summarized, as were absolute mean differences between the two levels of the variables and the 95% absolute confidence intervals for the absolute differences. Results were clearcut in that treatments were either strong or far from significant. Primary findings suggest that class of postage, incentives (especially when enclosed), followup, precontact, personalization of the cover letter, personalization of postage, and sponsorship to be effective methods of inducing responses to mail surveys. Questionnaire length also has an effect on response rate. Found to be ineffective were use of anonymity, color, cover letter appeals, deadlines, personalization of labels, and the status of the sender. Appendixes give category codings and correlations among dependent variables. (Contains 10 tables and 19 references.) (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 402 319

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

KATHY E. GREEN

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

Reviewing the Research on Mail Survey Response Rates: Meta-Analysis

Kathy E. Green

and

Susan R. Hutchinson

University of Denver

Presented as part of a symposium entitled "Reviewing the Research on Mail Survey Response Rates," SIG: Survey Research in Education, at the annual meeting of the American Educational Research Association, New York, April 1996.

TM025811

BEST COPY AVAILABLE

Introduction

Numerous methods have been proposed for cumulating results of research studies. These methods include the narrative review, counting statistically significant results, and various ways of averaging numerical outcome indices. Meta-analysis is one of the latter methods. The term, originated by Glass (1976), focuses on magnitudes of effect sizes.

Narrative review, also called literary or qualitative review, has the advantage of being useful when a small number of studies exist. Its limitations come in the subjectivity applied to integration of findings, in the failure to use all information available in studies, and in the severe difficulty of integrating findings of large numbers of studies. Cooper and Rosenthal (1980) found that even when the number of reviewed studies is as small as seven, a narrative review and a quantitative review will often reach different conclusions.

Glass' meta-analysis (Glass, McGaw, & Smith, 1981) emphasizes effect sizes and seeks to explain those effects with study characteristics (e.g., subject background variables, length of treatment). It uses more of the information available from individual studies and provides more accurate numerical estimates than narrative review, and has been shown to be robust when submited to reanalysis (Bangert-Drowns & Rudner, 1991). Regressions of effects on study characteristics, however, have generally yielded null findings. Applications of meta-analysis using Glass' methods have been critized on several counts. First, many studies produce more than one effect size and these are treated as independent though they are not, violating the assumption of independence required for most statistical tests. Hopkins and Gullickson (1992) assessed effects of enclosed and promised incentives on mail survey return rates for independent and nonindependent studies and found little difference between the two analyses. The effects of nonindependence may be more severe when the ratio of independent to nonindependent studies is low. Second, coding of independent variables often results in loss of information about qualitative differences in those variables, reducing the outcome of the analysis to gross rather than fine-grained statements. Third, these methods capitalize on chance, with potentially deceptive regression weights obtained when the number of studies is small.

Psychometric meta-analysis (Hunter & Schmidt, 1990) also focuses on effect sizes but corrects them for measurement error and other artifacts, providing an interval around an estimated true effect size. Use of this method presumes one has information such as instrument reliability and range restrictions. The psychometric meta-analyses of Hunter and Schmidt are an extension of Glass' methods rather than an incompatible alternative.

The present study used Glass' methods to empirically summarize the research on mail survey response rate manipulations. Boser and Clark (1996) conducted a narrative review using the same set of studies. The purpose of this presentation is two-fold. First, the results obtained speak to the similarities and differences one may see when analyzing common data with two different methods. Second, results provide a synthesis of literature regarding the effectiveness and strength of treatments used to enhance mail survey response rates as well as the predictive strength of selected characteristics of the survey setting.

Reviews of the Response Rate Literature

Early reviews of the effectiveness of treatments used to enhance mail survey response rates were narrative. This method was joined by regression approaches and then meta-analyses as ways of providing quantitative indices of the size of treatment effects and the significance of variables reflective of the survey setting. The results of selected recent reviews are briefly summarized here to provide a context for interpretation of the results of the present meta-analysis. For a more complete review of the reviews, consult Boser and Clark (1993) and Rodgers and Worthen (1995).

Anonymity. Jobber (1986) suggests anonymity is preferable to identification of survey respondents, due to a slight average improvement in response rates but also due to the potential for more accurate reporting of sensitive information. Yammarino, Skinner, and Childers (1991) found no significant effect for anonymity.

Color. Color of the survey paper has been found to have no substantial effect on response rates (Jobber, 1986), though a significant effect was found by Fox, Crask, and Kim (1988) with an average gain of 2% in responses when using green paper rather than white paper. Differences between treatment conditions for the 10 observations obtained from three studies ranged from -6% to +9%.

Cover Letter Appeals. Jobber (1986) found this treatment to have nonsignificant effects, though with a slight edge for a "help-the-sponsor" approach over social utility or egoistic appeals. Yammarino, Skinner, and Childers (1991) found significant response rate differences due to type of appeal, but did not define what they meant by an appeal. They found no effects of survey setting variables on the effect of appeals.

Deadlines. Use of deadlines has been supported due to their effects on response speed not on response rates (Jobber, 1986). Fox, Crask, and Kim (1988) found a mean effect of 1.7% of notifying survey recipients of a cutoff date. Study effects ranged from -14% to +8%. They identified 12 observations from a sample of 5 studies.

Follow-up. Rodgers and Worthen (1995) found an average gain of 14% across all forms of follow-up (letter, postcard, telephone) based on 15 observations. Differences between follow-up and no follow-up conditions ranged from 10% to 20%. Bruvold and Comer (1988) found follow-up contacts to have a stronger effect on responses than advance contacts.

Incentives. Two reviews of the effects of incentives on response rates were conducted using meta-analysis, one using Glassian methods on 85 observations (Hopkins & Gullickson, 1992) and the second using psychometric meta-analysis on 74 observations (Church, 1993). Hopkins and Gullickson found an average increase of 19% (95% CI: 17%, 22%) with enclosed incentives and an average 7% (95% CI: 3%, 12%) with promised incentives. Church found a 24% average increase for enclosed incentives. Both studies found a significant relationship between amount of the enclosed monetary gratuity and response rate. Neither review found any effect for context variables (population description, questionnaire length, and year of publication).

Personalization. Personalization manipulations may be applied to the cover letter and survey, to the mailing labels, or in the form of stamped envelopes versus metered envelopes. Jobber (1986) suggests personalization of the cover letter/survey to wield no power over response rates, and cautions that personalization may be counterproductive when sensitive information is

desired. In contrast, Rodgers and Worthen (1995) found a small (6.5%) effect of personalized versus generic cover letters.

Armstrong and Lusk (1987) located 6 studies differencing response rates for metered versus stamped envelopes. Stamps included regular, commemorative, and multiple small-denomination stamps. An average gain of 3% was found with use of greater personalization of postage, which was nonsignificant.

Postage Class. Armstrong and Lusk (1987) reviewed 34 studies of the differing effects of classes of return postage on response rates. In 20 studies listing differences in response rates between first class postage and business reply mail, an average gain of 9% (95% CI: 7%, 11%) was found. This is higher than the average gain reported by Fox, Crask, and Kim (1988) of 6%, with differences ranging from -10% to +32%.

Precontact. Effects of precontact on response rates have been reviewed by Schlegelmilch and Diamantopoulos (1991), Haggett and Mitchell (1994), and Rodgers and Worthen (1995), among others. Schlegelmilch and Diamantopoulos examined 27 publications with 73 observations, Haggett and Mitchell found 11 studies for 15 observations dealing strictly with industrial populations, and Rodgers and Worthen based their analysis on 25 observations. Schlegelmilch and Diamantopoulos found an average gain of 13% (differences ranged from -16% to +51%) over all forms of precontact, with an average 8% (differences ranged from -16% to +26%) for mail precontact. Haggett and Mitchell found an average gain of about 6% (from -7 to +16%) over mostly letter precontact. Rodgers and Worthen reported an average gain of 10% over all forms of precontact.

Questionnaire Length. Yammarino, Skinner, and Childers (1991) found response rates to questionnaires longer than four pages to be lower than those to shorter questionnaires. Heberlein and Baumgartner (1978), in contrast, found no significant effect of questionnaire length on response rates.

Sponsorship. Fox, Crask, and Kim (1988) found university sponsorship to be a strong response facilitator, with an average gain of 9% (study range from -11% to +23%). Yammarino, Skinner, and Childers (1991) also suggest a gain, but found it minimal.

When numerous studies exist in a field, and several attempts have been made to integrate their findings, it may be inevitable that conflicting results arise. Reviews of treatments affecting response rates have differed in variable definition, in study location and inclusion criteria, in method of study integration, and in complexity of design (i.e., simple effects of treatment versus interactive or confounded effects). The reviews have come to differing conclusions regarding the effect size of treatments designed to affect response rates. For example, Fox, Crask, and Kim (1988) found university sponsorship to be highly effective while Yammarino, Skinner, and Childers (1991) found it minimally effective. There is also disagreement about the confounding effects of variables describing the survey setting on response rates. Significant effects of setting variables have been found with regression approaches to study integration but less frequently with Glassian meta-analytic techniques. There is more consistency regarding whether major response facilitators such as incentives and follow-up are significant or not. A contribution of the present study is the opportunity it affords the reader to reflect on the differences in results found when the same set of studies were approached using two different integration methods.

Method

Appropriate articles were identified, selected, information abstracted, and the data coded. Article location criteria have been listed in the first

symposium paper as have article deletion criteria. We now move to a review of the criteria for including studies within articles in the analysis and then to the categories and procedures used to abstract and code data in this study.

Criteria for Inclusion

While a common set of articles was the starting point for this meta-analysis and the narrative review, both studies within articles, articles, and treatment variables were eliminated from consideration in the meta-analysis. Table 1 presents a list of studies and articles that were dropped since they were unclassifiable. These studies/articles were single instances of addressing a narrowly-defined topic and utilized treatments that with even the most liberal interpretation could not be included in a level of one of the defined treatment variables. Further, an arbitrary minimum of five articles was set for performing an analysis. Experimental treatments reported in fewer than five articles were excluded. As indicated in Table 1, data regarding salience were not analyzed nor were data regarding survey item variations or survey format manipulations, and we are mute on the effectiveness of those variables in inducing responses.

Table 1 here

Abstracting and Coding Studies

Since studies were to be divided into two sets and each set reviewed by one of the two authors, interrater agreement in information abstracting was assessed. First, the variables to be abstracted were identified. They are listed in Table 2 along with the percentage of the sample with available information for that variable. Studies were dropped from the meta-analysis if no information was available on experimental manipulation and the corresponding response rates.

Table 2 here

Both reviewers read a common set of 20 articles and reached the levels of agreement presented in Table 3. Discrepancies were:

- Sample Size: Rater 1 counted nondeliverables in the sample size and Rater 2 did not. Decision was to subtract nondeliverables from total sample size.
- Random Sample: Rater 2 counted systematic samples as random while Rater 1 did not. Decision was to treat systematic samples as random.
- Cover Letter: Rater 2 counted cover letters as part of treatment as "no" responses and Rater 1 as "yes." Decision was to count as "yes" responses.
- Follow-up: Rater 2 counted follow-up as "no" if treatment and Rater 1 did not. Decision was to treat as "yes" if treatment.

- Random Assignment to Treatment: Rater 2 counted as random assignment if cell n's seemed equal; Rater 1 counted as random only if explicitly stated. Decision to count as random assignment only if stated.
- Response Rate: Disagreements were found when values had to be hand-calculated; i.e., when not given directly in the table.

Table 3 here

Disagreements were resolved for the 20-article sample. Since agreement was low regarding the crucial variable, response rate, a second sample of 10 articles was selected. Agreement rates were higher for the most critical variables (Table 3) and were considered adequate to proceed with independent review. (One disagreement on survey topic was due to discrepant descriptions within the article.)

Coding categories were defined for primary analyses and are provided in Appendix A. As has been pointed out by critics of meta-analysis, category definitions are of necessity broad and disregardful of special case studies. The criticism could be leveled at the present review. For example, the offer of feedback was classified as a promised incentive.

Definition of Independent Variables, Dependent Variables, and Analyses

The groupings of studies with sufficient numbers for analysis yielded fourteen independent variables. Variables are listed in Appendix A with their original coding. However, due to small sample sizes for some codes, categories were collapsed for analysis. The resulting variables and their levels are presented in Table 4. While thorough explication of the nature of the treatments may be found in the original studies or in review papers, the treatments will be briefly described here.

- Anonymity was defined as whether the respondent knew the survey was identifiable or not. The conditions were cued by statements in the cover letter or by visible or unobtrusive coding on the survey, for example.
- Color was manipulated simply by copying the survey on white paper or on colored paper.
- Cover letter appeals emphasized how the good of society would be bettered by provision of the requested information (a social utility, science, or altruistic appeal), how the survey sponsor would be aided by the return of the survey (help the sponsor, sponsor appeal), or how the respondent may have been benefited by return of the survey (egoistic appeal), for example, by better programming on television.
- Deadlines included a due date in the cover letter or merely requested survey return.
- Follow-ups were phone calls, letters, postcards, or second copies of the survey sent to nonrespondents to the original mailing.
- Incentives were enclosed tokens (e.g., tea bags, pens, money, checks) or promised tokens (e.g., money, copy of results of the survey, gift).

- Personalization of the cover letter/survey was accomplished by hand-signing cover letters, personal inside addresses instead of Dear Ms/Sir, 'user-friendliness,' use of nicknames on the cover letter, computer or hand-typed instead of mimeographed letters, handwritten postscripts, and combination of these elements and variations intended to make the respondent feel the survey was meant for him/her individually. Mailing labels were personalized if they appeared hand-typed rather than machine-generated; postage was personalized if, for example, regular or commemorative stamps were used rather than a business reply envelope.
- Questionnaire length was assessed by number of pages or number of items.
- Postage class reflected mailing costs with variations being bulk mail, first class, air mail, special delivery, or certified mail.
- Precontact employed phone, postcard, or letter to notify the survey recipients that they would be receiving an important survey.
- Survey sponsors included universities, research firms, commercial organizations, or charities. Sponsorship was indicated by the letterhead.
- Status of the cover letter signator was varied as higher (e.g., a professor, CEO) or lower (e.g., a student, office manager).

Table 4 here

Five dependent variables were defined, all variations on the theme of differences in response rates. They were:

- The first dependent variable was formed by treating the percentage of responses for the control or comparison group and the percentage of responses for the experimental group as a within-study variable in a repeated measures analysis of variance.

- The second dependent variable, the absolute difference, was the difference between the response rate for one experimental condition and the response rate for a second experimental or comparison condition. Treatments such as cover letter appeals had multiple conditions with none defined as a control condition.

- The third dependent variable was the relative increase in returns calculated as the ratio of the response rate for the experimental group to the response rate for the comparison group.

- As calculation of means and standard deviations for categorical variables may be statistically problematic, Glass et al. (1981) suggest a means of translating proportions based on polytomies into a hypothetical metric variable. The rationale for doing so is the same as that underlying probit analysis (Finney, 1971). The method is to difference the standard normal deviates corresponding to the percentages observed in the experimental and comparison groups. (See Glass et al., pp. 136-138, for details.) This difference formed the fourth dependent variable.

- The fifth and last dependent variable was formed by differencing the log-odds of the response rate for the treatment and comparison groups. Log-odds or logits were calculated as: $\ln(p/1-p)$.

Primary analyses were conducted to determine the effect of each of the 14 treatment variables on response rate using repeated measures analysis of variance (ANOVA). Although ANOVA is designed for use with interval or ratio measures of dependent variables, Harris (1994) finds the use of ANOVA with proportions feasible unless extreme proportions (close to 0.0 or 1.0) or a very close decision on an issue warrants a more sophisticated approach. Repeated measures ANOVA further requires a positive correlation between levels of the repeated measures for effective use. The correlations between groups defined by level of treatment, presented in Table 5, were positive and significant for all of the 14 independent variables. Subsequent analyses examined the separate influence of random assignment to treatment, topic (targeted or not), random sample or not, population description (general versus other), and journal type (business versus other) using repeated measure analyses of variance. Not all coded variables were selected for subsequent analyses due to lack of clear information for that variable or lack of variation in practice, e.g., surveys are most often accompanied by cover letters and SASEs.

Table 5 here

Prior to analyses, the probit and logit transformations, relative increase, and the absolute difference dependent variables were correlated. As correlations were high (Appendix B), the conceptually transparent variable of absolute difference (AD) was used in preference to the probit, logit, and relative increase transformations. The distribution of the effects was described using the mean difference and 95% confidence interval. Joint effects of survey setting characteristics were examined using multiple regressions with the absolute difference as the dependent variable.

Results

Table 6 presents a summary of the significance of the repeated measure main effects (response rate) across experimental and comparison groups for the treatment conditions. Table 6 also presents the mean absolute difference between the two levels of the variables and the 95% confidence interval for the absolute difference. For example, the mean difference in response rates between surveys with and without a deadline was 1.11%, with a confidence interval around the mean difference of -3.6% to +5.9%. Results were clearcut. Treatments were either strong or far from significant, with the single exception of questionnaire length where the effect was marginal. The direction of effects was consistent with expectation. For sponsorship, where the literature has sometimes suggested a minimal benefit of university over other sponsors, the present analysis suggests a clear benefit of university over other sponsors. Table 7 presents the mean response rate by level for each of the 14 treatment variables.

Tables 6 and 7 here

The effects of five characteristics of the survey setting on differences between levels of treatment in response rates in concert with each of the 14 treatment variables were examined. Numerous combinations of survey setting characteristics and independent variables could not be analyzed due to the small sample sizes and thus cells with zero entries. Significant interactions between treatment and setting variables and main effects of setting variables are listed in Table 8. There were few relationships of any magnitude between setting characteristics and treatment variables.

Table 8

The nature of the effects found for survey setting variables are clarified by review of the cell means provided in Table 9. The significant interaction of cover letter and random assignment is suspect since one cell (no random assignment, sponsor appeal) had but one study representing that condition. The only other interaction of consequence was between follow-ups and the population description. Follow-up was effective with both general populations and professional populations but increased response rates more sharply when the population was professional rather than general.

Table 9

Multiple regressions incorporated all five survey setting characteristics in prediction of absolute differences in the response rates for the fourteen treatment conditions. As can be seen from Table 10, although the multiple correlation coefficients were high in many cases, there were only two treatments for which the survey setting variables provided statistically significant prediction, cover letter appeal (social utility versus sponsor) and follow-up. Low statistical power resulting from small sample sizes explains the lack of significance.

Table 10

Discussion

The primary finding of this study suggests class of postage, incentives (especially enclosed), follow-up, precontact, personalization of cover letter/survey, personalization of postage, and sponsorship to be effective methods of inducing responses to mail surveys. Questionnaire length has an effect on response rate as well. Found to be ineffective were use of anonymity, color, cover letter appeals, deadlines, personalization of labels, and the status of the sender. Results of this meta-analysis are generally quite similar to those found by previous researchers.

Previous meta-analyses of incentives (Church, 1993; Hopkins & Gullickson, 1992) found a higher effect for enclosed monetary incentives than was found in this study for any type of enclosed incentive (24% and 18% versus

15.49%). Mean absolute differences for follow-up, precontact, personalization, color, deadlines, and sponsorship were all similar to results found by Rodgers and Worthen (1995), Fox, Crask and Kim (1988), and Armstrong and Lusk (1987). The mean absolute difference for questionnaire length was slightly higher than that reported by Rodgers and Worthen. The most striking difference in this study was a higher difference in response rates due to class of postage than that found by Armstrong and Lusk. Another noteworthy finding is agreement with Fox, Crask, and Kim regarding the effect of sponsorship. They also found a 9% increase, in contrast with the minimal effect found by Yammarino, Skinner, and Childers (1991).

The characteristics of the survey setting had little relationship to the effectiveness of treatment variables. Such null results have been noted in meta-analyses in other fields, which also found little effects of contextual variables. Effects of context may exist, but for most of the treatments described in this study, the sample of articles was too small to reliably detect them.

Multiple dependent variables were defined in this study. Transformations of the absolute difference were all highly correlated, a result consistent with that of Hopkins and Gullickson (1992). While transformed variables may possess preferable statistical characteristics, their interpretation is clouded by their complexity. Further development of the statistical analysis of proportions would be desirable.

Basic information was missing from many of the articles reviewed. Over half the articles had no indication of the survey length and lacked information about the procedure. Sampling methods were unclear and authors were unclear about whether the study used random assignment. A number of articles were discarded because response rates were not numerically presented. Reporting standards for survey studies should be developed and adhered to, including reporting of response rates for nonsignificant results. Provision of basic information would allow reviewers to eventually compile a better framework for understanding response rates to mail surveys.

Care should be taken by the survey researcher in projecting the likely result of employing a response rate manipulation. As can be seen from the range of differences between treatment and comparison conditions in Table 6, there was but one treatment (sponsorship) for which differences were positive in all cases. Even enclosed incentives served to depress the response rate for at least one unlucky survey researcher. This study describes aggregate effects of small numbers of studies, not natural laws.

References

Armstrong, J. S., & Lusk, E. J. (1987). Return postage in mail surveys: A meta-analysis. Public Opinion Quarterly, 51, 233-248.

Bangert-Drowns, R. L., & Rudner, L. M. (1991). Meta-analysis in Educational Research. Wahington, DC: ERIC Clearinghouse on Tests, Measurement, and Evaluation, EDO-TM-91-11.

Boser, J. A., & Clark, S. B. (1996, April). Reviewing the research on mail survey response rates: Narrative review. Paper presented at the annual metting of the American Educational Research Association, New York.

Boser, J. A., & Clark, S. B. (1993, April). Response rates in mail surveys: A review of the reviews. Paper presented at the annual metting of the American Educational Research Association, Atlanta.

Bruvold, N. T., & Comer, J. M. (1988). A model for estimating the response rate to a mailed survey. Journal of Business Research, 16, 101-116.

Bruvold, N. T., Comer, J. M., & Rospert, A. M. (1990). Interactive effects of major response facilitators. Decision Sciences, 21, 551-562.

Church, A. H. (1993). Estimating the effect of incentives on mail survey response rates: A meta-analysis. Public Opinion Quarterly, 57, 62-79.

Cooper, H. M., & Rosenthal, R. (1980). Statistical versus traditional procedures for summarizing research findings. Psychological Bulletin, 87, 442-449.

Finney, D. J. (1971). Probit Analysis. Cambridge, England: Cambridge University Press.

Fox, R. J., Crask, M. R., & Kim, J. (1988). Mail survey response rate: A meta-analysis of selected techniques for inducing response. Public Opinion Quarterly, 52, 467-491.

Glass, G. V. (1976). Primary, secondary, and meta-analysis of research. Educational Researcher, 5, 3-8.

Glass, G. V., McGaw, B., & Smith, M. L. (1981). Meta-analysis in Social Research. Beverly Hills, CA: Sage.

Harris, R. J. (1994). ANNOVA: An analysis of variance primer. Itasca, IL: F. E. Peacock.

Heberlein, T. A., & Baumgartner, R. (1978). Factors affecting responses to mailed questionnaires: A quantitative analysis of the published literature. American Sociological Review, 43, 447-462.

Hopkins, K. D., & Gullickson, A. R. (1992. Response rates in survey research: A meta-analysis of the effects of monetary gratuities. Journal of Experimental Education, 6, 52-62.

Hunter, J. E., & Schmidt, F. L. (1990). Methods of Meta-analysis. Newbury Park, CA: Sage.

Jobber, D. (1986). Improving response rates in industrial mail surveys. Industrial Marketng Management, 15, 183-195.

Rodgers, P., & Worthen, B. R. (1995, April). A meta-analysis of factors that influence the response rate of mailed questionnaires. Paper presented at the annual meeting of the American Educational Research Association, San Francisco.

Yammarino, F. J., Skinner, S. J., & Childers, T. L. (1991). Understanding mail survey response behavior: A meta-analysis. Public Opinion Quarterly, 55, 613-639.

Table 1

Articles, Studies, and Treatments Deleted from Analysis

Author/Date	Topic	Reason for Exclusion
Armstrong, 1991	Type of return address	Only paper on topic
Boser, 1990a	Survey booklet format	Only paper on topic
Brown, 1965	Screening questions	Only paper on topic
Childers & Ferrell, 1979	Size of paper	Only paper on topic
Dillman et al., 1993	Type of survey form	Only paper on topic
Dommeyer et al., 1991	Teaser on envelope	Only paper on topic
Dommeyer, 1987	Negative followup appeal	Only paper on topic
Enger et al., 1992	Survey format	Only paper on topic
Field, 1979	Gender of investigator	Only paper on topic
Ferrell et al., 1984	Where receive survey	2 in set
Ferriss, 1951	SASE or not	2 in set
Ford, 1968	Q format	Only paper on topic
Frey, 1991	Item order, cover	Only paper on topic
Friedman & Goldstein, 1975	Researcher ethnicity	Only paper on topic
Goldstein & Friedman, 1975	Cover letter form	Only paper on topic
Goulet, 1977	Follow-up letter	Only paper on topic
Hammond, 1959	Cover letter content	Only paper on topic
Jacobs, 1986	Response format on Q	Only paper on topic
Johnson et al., 1993	Q size	Only paper on topic
Jones & Lang, 1980	Q density	Only paper on topic
Kerin & Harvey, 1976	SASE stamped vs. not	Only paper on topic
Kimball, 1961	Permit vs BR envelope	Only paper on topic
Labreque, 1978	Owner vs. manager sponsored	Only paper on topic
Martin, 1970/71	Salience	<5 in set
McGinnis & Hollon, 1977	Where receive survey	2 in set
Peterson et al., 1989	Salience	<5 in set
Pressley & Tullar, 1977	Cartoons on survey	Only paper on topic
Rucker et al., 1984	Salience	< 5 in set
Rudd & Maxwell, 1980	Salience	< 5 in set
Senf, 1987	Reason for no response	Only paper on topic
Sheth et al., 1980	Q questions	Only paper on topic
Sirken et al., 1960	Cover letter tone	Only paper on topic
Smith & Bers, 1987	Number of follow-ups	Only paper on topic
Stevens, 1974	Q computer coded	Only paper on topic
Swan & Epley, 1981	Income question	Only paper on topic
Taylor et al., 1987	Familiarity with sponsor	Only paper on topic
Tollefson et al., 1984	Q appearance	Only paper on topic
Wagner & O'Toole, 1986	Use of humor	Only paper on topic

Table 2

Variables Abstracted and Percent of Studies with Information Available

Variable	% Available	% Unclear/Unspecified
Author	100%	
Date	100%	
Journal	100%	
Population Description	99%	
Sample Size	100%	
Use of Random Sample or Not	80%	20%
Survey Topic	82%	18%
Survey Length in Pages	43%	57%
Survey Length in Items	33%	68%
Survey Accompanied by Cover Letter or Not	78%	22%
Survey Accompanied by SASE or Not	50%	50%
Follow-Ups or Not	97%	3%
Random Assignment to Treatment	77%	23%
Study Clarity	100%	Researcher Defined

Table 3

Interrater Response Agreement

Variable	Sample 1 (n=20) % Agreement	Sample 2 (n=10) % Agreement
Population Description	95%	100%
Sample Size	65%	88%
Random Sample	85%	100%
Survey Topic	100%	75%
Survey Length	100%	100%
Survey Accompanied by Cover Letter	90%	88%
Survey Accompanied by SASE	90%	88%
Follow-ups Used	90%	100%
Random Assignment to Treatment	85%	100%
Treatments	100%	100%
Response Rates	80%	100%

Table 4

Independent Variables and Categories Used in Analyses

Variable	Category 1	Category 2	Category 3
Anonymity	Yes	No	
Color	Color	White	
Cover Letter Appeals	Social Utility	Help the Sponsor	Egoistic
Deadlines	Yes	No	
Followups	Yes	No	
Incentives	Enclosed	Promised	None
Personalization:			
Cover Letter/Survey	More	Less	
Labels	More	Less	
Postage	More	Less	
Postage Class	First Class	Higher than First Class (certified, special delivery)	
Precontact	Yes	No	
Questionnaire Length	Longer	Shorter	
Sponsorship	University	Commercial/Research Firm	
Status of Sender	More	Less	

Table 5

Correlations between Levels of Treatment Variable Response Rates

Variable	Correlations
Anonymity	.96***
Color	.96***
Cover Letter Appeals[a]	.86 to .92**
Deadlines	.94**
Followups	.92***
Incentives[a]	.85 to .96***
Personalization:	
Cover Letter/Survey	.92***
Label	.97***
Postage	.89***
Postage Class	.78***
Precontact	.86***
Questionnaire Length	.85***
Sponsorship	.83***
Status of Sender	.99***

[a]Multiple levels of treatment variable therefore multiple correlations among levels were calculated.

**p<.01 ***p<.001

Table 6

Effects of Treatment on Survey Response Rate

Treatment	F	df	p	Mean AD	95% CI	Range of AD
Anonymity	.18	1,12	.68	.70	(-2.9, 4.3)	-7.2 to 13.4
Color (vs. White)	.62	1,6	.46	1.26	(-2.6, 5.2)	-6.4 to 6.9
Cover Letter Appeals	3.37	1,7	.11			
Social Utility-Sponsor				-2.25	(-7.1, 2.6)	-17.5 to 4.7
Sponsor-Egoistic				-4.14	(-9.5, 1.2)	-14.0 to 3.0
Social Utility-Egoistic				-1.98	(-10.2, 6.3)	-21.4 to 11.0
Deadlines	.33	1,6	.59	1.11	(-3.6, 5.9)	-5.8 to 7.6
Followup (vs. None)	29.59	1,12	.00	12.30	(7.4, 17.2)	-.9 to 27.3
Incentive:						
Enclosed, Promised, None	36.80	2,32	.001	10.77	(8.7, 12.8)	-7.0 to 32.5
Enclosed vs. None	46.22	1,51	.001	15.49	(12.8, 18.2)	-7.0 to 38.4
Promised vs. None	12.47	1,31	.001	3.00	(1.3, 4.7)	-6.5 to 10.5
Enclosed vs. Promised	32.91	1,17	.001	15.60	(9.9, 21.3)	-5.2 to 44.9
Personalization:						
Cover Letter/Survey	15.21	1,45	.001	3.68	(1.8, 5.6)	-15.5 to 15.0
Labels	.34	1,9	.57	.95	(-2.7, 4.6)	-7.0 to 11.0
Postage	17.35	1,27	.001	5.72	(2.9, 8.5)	-3.8 to 31.0
Postage Class	32.28	1,10	.001	16.08	(9.8, 22.4)	-3.0 to 27.5
Precontact	28.33	1,32	.001	10.40	(6.6, 14.2)	-11.0 to 37.0
Questionnaire Length	4.91	1,18	.04	4.69	(.2, 9.1)	-11.9 to 22.0
Sponsorship	21.16	1,6	.004	9.45	(4.4, 14.5)	1.9 to 19.1
Status of Sender	.03	1.6	.87	.16	(-2.0, 2.3)	-3.0 to 3.2

Note: F = test of repeated measures ANOVA for treatment main effects with associated df and p value; AD = absolute difference in response rates between treatment and comparison groups.

Table 7

Response Rates for Treatment Variables by Level

Treatment	Level 1	Mean	SD	Level 2	Mean	SD	Level 3	Mean	SD
Anonymity	Yes	52%	21%	No	51%	21%			
Color	Color	33%	15%	White	32%	15%			
Cover Letter Appeals	Social	40%	12%	Sponsor	40%	14%	Egoistic	43%	17%
Deadlines	Yes	35%	10%	No	34%	13%			
Followup	Yes	46%	20%	No	35%	14%			
Incentive	Enclose	49%	18%	Promise	33%	19%	None	32%	18%
Personalization:									
Cover Letter, Survey	More	41%	17%	Less	38%	16%			
Labels	More	40%	20%	Less	39%	21%			
Postage	More	39%	16%	Less	33%	13%			
Postage Class	First	27%	12%	Higher	43%	15%			
Precontact	Yes	48%	17%	No	38%	22%			
Questionnaire Length	Longer	45%	17%	Shorter	50%	17%			
Sponsorship	Univ	40%	8%	Other	30%	10%			
Status of Signator	Higher	47%	18%	Lower	47%	18%			

Table 8

Effects of Characteristics of the Survey Setting on Response Rates

Treatment Variable	Survey Setting	F	df	p
Cover Letter Appeal	RA to Treatment	14.16	1,8	.006
Social Utility vs. Help the Sponsor Interaction[a]				
Followup (vs. None)				
Main Effect	RA to Treatment	6.50	1,11	.03
Interaction	Topic	6.06	1,11	.04
Interaction	Population	11.47	1,11	.006
Incentive:				
Enclosed vs. Promised	Population	4.30	1,49	.05
Interaction				
Enclosed vs. None	Population	4.95	1,16	.05
Main Effect				
Personalization:				
Cover Letter, Survey Interaction	Journal	5.22	1,44	.03
Postage Class Interaction	Random Sample	6.38	1,9	.04

Note: Dependent variable in ANOVAs was difference in response rate between levels of treatment.

[a]Two of four cells with n = 1.

Table 9

Cell Mean Response Rates for Main and Interactive Effects of Setting Variables

Effect	Setting Variable	Treatment Variable Level 1	Level 2
		Cover Letter Appeals	
Interaction	Random Assignment:	Social	Sponsor
	Yes	39%	47%
	No	39%	64%
		Follow-up	
Interaction	Topic:	Yes	No
	Targeted	56%	38%
	Nontargeted	42%	33%
		Follow-up	
Interaction	Population:	Yes	No
	General	41%	34%
	Other	54%	36%
		Postage	
Interaction	Random Sample:	First	Higher
	Yes	33%	40%
	No	25%	45%
		Personalization (Letter, Survey)	
Interaction	Journal:	More	Less
	Business	39%	37%
	Other	44%	38%
		Incentive	
Interaction	Population:	Enclosed	Promised
	General	38%	23%
	Other	57%	42%
		Follow-up	
Main Effect	Random Assignment:		
	Yes	36%	
	No	58%	
		Enclosed/Promised Incentives	
Main Effect	Population:		
	General	36%	
	Other	45%	

Table 10

Regression of Response Rate on Treatment and Survey Setting Characteristics

Treatment	R	F	p	Significant Predictors
Anonymity	.61	.83	.56	
Color	.96	5.53	.16	
Cover Letter Appeals:				
Social Utility - Sponsor	.97	13.33	.02	RA, RS, Journal
Sponsor - Egoistic	.81	1.46	.39	
Social Utility - Egoistic	.95	7.73	.07	
Deadlines	.63	.64	.64	
Followup	.88	4.97	.03	
Incentive:				
Incentive vs. None	.22	.64	.67	
Enclosed vs. None	.30	.88	.50	
Promised vs. None	.23	.28	.92	
Enclosed vs. Promised	.64	1.65	.22	
Personalization:				
Cover Letter, Survey	.48	2.37	.06	
Labels	.91	2.50	.41	
Postage	.35	.63	.68	
Postage Class	.90	4.49	.06	
Precontact	.31	.58	.71	
Questionnaire Length	.39	.47	.79	
Sponsorship	.91	1.54	.52	
Status of Sender	.71	1.05	.49	

Note: RA = random assignment to treatment or not; RS = randomly selected sample or not; Journal = business journal or journal from other discipline.

APPENDIX A

Category Coding

<u>Survey Context Variables</u>	
Author:	
Date:	
Journal:	1=business 2=education 3=psych 4=other
Population Description:	1=phone directory or other general public 2=profesional except teachers 3=college students/teachers 4=military 5=other 6=no information
Sample Size	
Random Sample:	0=no 1=yes 2=unclear 3=nonrespondents
Topic:	0=not specified 1=targeted 2=not targeted
Survey Length in Pages:	88=no info 99=treatment
Survey Length in Items:	26-28 888=no info 999=treatment
Cover Letter:	0=no 1=yes 2=unclear 3=treatment
SASE:	0=no 1=yes 2=unclear 3=treatment
Follow-Up:	0=no 1=yes 2=unclear 3=treatment
Random Assignment to Treatment:	0=no 1=yes 2=unclear
Clarity:	1=yes 2=moderate 3=no
<u>Treatments</u>	
Incentives:	Enclosed Promised None
Precontact:	Yes No
Personalization of cover letter or survey:	More Less
Labels personalized:	More Less
Postage personalized:	More (e.g., stamps) Less (e.g., metered, busines reply)
Status of sender:	More (e.g., professor) Less (e.g., grad student)
Questionnaire length:	Longer Shorter
Appeals:	Social Utility Help the Sponsor, Need Reply Egoistic, Benefit to Receiver
Sponsor:	University Commerical Firm Research Firm
Deadline:	Yes No
Anonymity:	Yes No
Type of postage:	First class Better than first class
Follow-up:	Phone Letter Postcard Other No
Color:	Color White
Salience:	More Less

APPENDIX B
Correlations Among Dependent Variables

Variable	AD	Probit	Logit
Anonymity			
Probit	.99		
Logit	.99	.99	
Relative Increase	-.90	-.94	-.94
Color			
Probit	.95		
Logit	.99	.96	
Relative Increase	.94	.94	.98
Cover Letter Appeals[a]			
Probit	.99		
Logit	.99	.99	
Deadlines			
Probit	.99		
Logit	.99	.99	
Relative Increase	.96	.98	.99
Follow-up			
Probit	.92		
Logit	.97	.96	
Relative Increase	.71	.78	.83
Incentives			
Probit	.90		
Logit	.88	.97	
Relative Increase	.60	.81	.87
Personalization of:			
Cover Letter/Survey			
Probit	.99		
Logit	.98	.99	
Relative Increase	.87	.98	.98
Label[a]			
Probit	.99		
Logit	.98	.99	
Postage			
Probit	.99		
Logit	.98	.99	
Relative Increase	.80	.83	.83
Postage Class			
Probit	-.95		
Logit	.93	-.99	
Relative Increase	.66	-.85	.88
Precontact			
Probit	.98		
Logit	.98	.99	
Relative Increase	.93	.96	.96
Questionnaire Length[a]			
Probit	-.99		
Logit	-.99	.99	

Variable	AD	Probit	Logit
Sponsorship[d]			
Probit	.87		
Logit	.99	.92	
Status of Sender			
Probit	.99		
Logit	.99	.99	
Relative Increase	.97	.99	.95

Note: AD = absolute difference.
[d]Relative increase not computed; lacks clear control group.